

Why people believe in indeterminist free will

Oisín Deery

Published online: 12 October 2014
© Springer Science+Business Media Dordrecht 2014

Abstract Recent empirical evidence indicates that (1) people tend to believe that they possess indeterminist free will, and (2) people’s experience of choosing and deciding is that they possess such freedom. Some also maintain that (3) people’s belief in indeterminist free will has its source in their experience of choosing and deciding. Yet there seem to be good reasons to resist endorsing (3). Despite this, I maintain that belief in indeterminist free will really does have its source in experience. I explain how this is so by appeal to the phenomenon of prospection, which is the mental simulation of future possibilities for the purpose of guiding action. Crucially, prospection can be experienced. And because of the way in which prospection models choice, it is easy for agents to experience and to believe that their choice is indeterministic. Yet this belief is not justified; the experience of prospection, and hence of free will as being indeterminist, is actually consistent with determinism.

Keywords Free will · Alternative possibilities · Determinism/indeterminism · Prospection · Causal modeling

1 Folk beliefs about indeterminism

People often judge free choice as incompatible with determinism—that is, as indeterministic. Determinism is the thesis that, given the actual past and laws of nature, there is at any time exactly one physically possible future. Many people also believe that they possess indeterminist free will: they believe that the nature of actual free decisions implies that determinism is false. Libertarianism is the

O. Deery (✉)
Department of Philosophy, University of Arizona, 213 Social Sciences, 1145 E. South Campus
Drive, Tucson, AZ 85721, USA
e-mail: oisín@oisindeery.com

philosophical formalization of this view: free will is inconsistent with determinism, and we are free. Although libertarianism is a minority view among philosophers, it is strongly implicated in ordinary thinking. Many people believe that the world is, as John M. Fischer puts it,¹ a “Garden of Forking Paths”:

We naturally think of the future as open. We think of the future as containing various paths that branch off one past; although we know we will travel along just one of these paths, we take it that some of the other paths are (at least sometimes) genuinely accessible to us. In deliberating and deciding on a course of action, we intuitively think of ourselves (at least sometimes) as determining which path to take, among various paths we *could* take. (Fischer 1994, p. 190)

Of course, compatibilists—who maintain that free will is compatible with determinism—often argue that, even assuming determinism, agents might have access to alternatives for action (List 2014; Ismael 2013; Fara 2008; Vihvelin 2004). Yet a number of recent studies indicate that *folk* thinking, at least, tends to regard alternatives for choice as indeterministic (Knobe 2014; Rose and Nichols 2013; Nichols and Knobe 2007; Nichols 2004). In one classic study (Nichols and Knobe 2007), participants were given descriptions of a deterministic universe (A) and another universe (B) in which everything is determined *except* human choices. Participants were asked “Which of these universes is most like ours?” More than 90 % of respondents said that universe (B)—the indeterministic universe—is most like our own (2007, p. 669). Further evidence supporting the view that people think of alternatives for choice as indeterministic comes from studies conducted by Nichols (2004). In one experiment, children were asked whether an agent was free to choose differently than she did, even if everything stayed the same right up until the moment she made her choice. A sample question went as follows:

Joan is in an ice cream store and wants some ice cream. She chooses vanilla
 ...If everything in the world was the same right up until she chose vanilla, did
 Joan have to choose vanilla? (2004, pp. 486–487)

Participants in this study were also given a contrast case involving a pot of boiling water, and were asked whether, if everything stayed the same right up until the water boiled, the water had to boil. The results showed a significant difference between responses to the agentive case and the case involving the boiling water. Participants tended to agree in the agentive case that the agent was free to choose differently, yet they tended to deny that the water could not have boiled. Nichols interprets these results as indicating that people tend to regard human choice as indeterministic.

Various other experimental studies challenge these results (Murray and Nahmias 2014; Nahmias and Murray 2011; Nahmias et al. 2007; Turner and Nahmias 2006). Yet despite this, there is converging evidence that ordinary thinking is at least *partly* indeterminist (May 2014; Deery et al. 2014; Feltz et al. 2012; Weigel 2011; Feltz

¹ Fischer borrows this phrase from Jorge Luis Borges.

and Cokely 2009). For compatibilists, it is a perplexing question how to explain the source of the widespread folk tendency to believe in indeterminist freedom.

2 Do indeterminist beliefs have their source in experience?

Where might the belief in indeterminism come from? Libertarians often point to experience as the source (e.g., Searle 1984; Campbell 1951). These philosophers claim that people's experience of deliberating and choosing leads them to believe that choice is indeterministic. Some even appeal to such experiences as evidence that choice *is* indeterministic (O'Connor 1995, pp. 196–197). In any case, the libertarian hypothesis is that agents experience the future as open in a way that it could not be if determinism is true.

While early studies by Nahmias et al. (2004) deny the libertarian hypothesis, more comprehensive studies conducted by Deery et al. (2013) indicate that the hypothesis may in fact be right. In these experiments, participants were asked to decide between two options (for instance, two charities) and were asked whether, as they faced their decision, they experienced being free to choose either option. Participants were allowed to interpret such freedom however they wished. Most participants reported experiencing being free to choose either way. Determinism was then explained to participants. Following comprehension checks, participants were asked whether the experience they had earlier reported was consistent with determinism. Most participants tended to judge their experience as inconsistent, whether the decision was present-focused or retrospective, imagined or actual, or morally salient or morally neutral. This suggests that libertarians may be correct in claiming that the belief in indeterminist freedom has its source in experience.

Yet there is reason to resist this explanatory move. According to Nichols (2012), it is implausible that belief in indeterminist freedom has its source in experience. Nichols concedes that some beliefs—such as the belief that it is currently raining outside—might come directly from raw experience. But indeterminism is too sophisticated a notion to grasp in this way: "...as an explanation for the belief in indeterminism, the appeal to experience is too anemic to be convincing" (2012, p. 293).

This paper makes the case that experience is indeed full-blooded enough to be judged as indeterminist, and that people's belief in indeterminism has its source in their ordinary experience of deliberation and choice. Yet the account I offer is deflationary, in the sense that while it explains the belief in indeterminist freedom as having its source in experience, it nonetheless says that this belief is not justified; hence, it is consistent with the truth of determinism and does not provide evidence for libertarianism.

3 Etiology of indeterminist experience and belief

Belief in indeterminist free will has its source in the experience of prospection,² which is the mental simulation of future possibilities for the purpose of guiding

² The term 'prospection' is due originally to Gilbert and Wilson (2007) and Buckner and Carroll (2007).

action. As I explain shortly, these simulations function as competitors to perceptual experience, in the sense that they can similarly affect an agent's feelings and motivation, and in such a way that the simulations count as genuinely experiential.

Prospection explains how agents can experience mere possibilities, as I explain in a moment. By itself, however, prospection cannot explain how an agent might experience indeterministic possibilities, or her decisions as undetermined by her own prior mental states. By contrast, causal modeling explains how a deliberating agent's cognition might well treat her available alternative possibilities as indeterministically open. Yet causal modeling by itself fails to indicate a mechanism by which an agent might arrive at an indeterministic belief as a result of such an experience.

The account I develop puts these two components together, in such a way that each component compensates for the shortcomings of the other. When the experience of prospection is cashed out in terms of causal modeling, the result is a mechanism by which people arrive at indeterminist beliefs about free will.

3.1 Prospection is forward-looking

According to Gilbert and Wilson (2007), "Just as retrospection refers to our ability to experience the past, prospection refers to our ability to 'pre-experience' the future by simulating it in our minds" (2007, p. 1352). For Gilbert and Wilson, prospection promotes our survival by enabling us to predict future hedonic reward in a sophisticated way. "As the mere thought of eating a liver popsicle reveals," Gilbert and Wilson write, "mental simulations of the future can elicit hedonic reactions in the present" (2007, p. 1352).³

Yet for Gilbert and Wilson, prospection is episodic, and its simulations are "cardboard cut-outs of reality," barely convincing enough to "elicit brief hedonic reactions" (2007, p. 1354). On this view, prospection "cannot generate simulations that have all the richness and reality of genuine perceptions" (2007, p. 1354).

By contrast, on a more sophisticated view developed by Seligman et al. (2013), prospection occurs continuously and is far more powerful than on Gilbert and Wilson's model. According to Seligman et al.,

Prospection... is guidance... by present, evaluative representations of possible future states. These representations can be understood minimally as 'If X, then Y' conditionals, and the process of prospection can be understood as the generation and evaluation of these conditionals (2013, p. 119).⁴

Considered within this framework,

³ Nonetheless, as Gilbert and Wilson (2007) note, prospection may be prone to a number of errors.

⁴ It is unclear whether Seligman et al. think *all* counterfactual cognition can be explained in terms of prospection. At any rate, I do not endorse this claim. After all, the predictive coding at work in many accounts of prospection, since it occurs at the implementation level in the brain, is presumably too low-level to be capable of explaining all conscious counterfactual reasoning. More plausibly, conscious prospection is a special case of counterfactual reasoning, and the best way to understand this sort of reasoning is, I maintain, in terms of interventionist causal modeling.

[P]eople and intelligent animals draw on experience to update a branching array of evaluative prospects that fan out before them. Action is then selected in light of their needs and goals. ...These prospects can include not only possibilities that have occurred before but also possibilities that have never occurred—and these new possibilities often play a decisive role in the selection of action. (2013, p. 119)

For Seligman et al., agents—in order to regulate their interactions with the environment—construct representational mental models of that environment. The most efficient models will be of the form, “*if* in circumstance C and state S, *then* behavior B has outcome O with probability p” (2013, p. 124). Typically, these “feedforward/feedback” models will have the following structure:

expectation → observation → discrepancy detection → discrepancy-reducing change in expectation → expectation → ...

As Seligman et al. put it,

Expectation is pivotal in [this] schema... because it transforms experience into experimentation—continuously generating a “test probe” so that the next experience always involves an implicit question and supplies an answer, which can then function as an error-reducing “learning signal.” (2013, p. 124)

In this way, agents generate and use simulations of future possibilities, often by drawing on past experience, and the function of these simulations is to enable the agents to navigate effectively into the future by selecting appropriate actions.⁵ For Seligman et al., such a model makes human agency intelligible:

The driven-by-the-past framework makes agency and choice difficult to understand—individuals are responders rather than navigators. ...If instead we see the individual as using past experience as information, as continually forming and evaluating a range of future possibilities, and as electing action from among these possibilities in light of what she likes and values, then we can see that active agency is a natural part of the causal structure of action. (2013, p. 127)

⁵ Seligman et al. concede that many actions do not require prospection. When actions “can be successfully repeated without need for evaluation of alternatives” (2013, p. 125), they move from being under *teleological* control [selection “in light of values and goals” (p. 120)] to being under *habitual* control. Moreover, while Seligman et al. maintain that motivation is often teleological, they concede that drive-like, non-goal-directed motivation *sometimes* occurs: “Addiction and salt deprivation, for example, can produce wanting without liking... Certain physiological demands, natural or artificial, can produce ‘driven’ motivation even in the face of profound distaste and resistance, but this is atypical indeed. Ordinary action, even eating a meal when hungry, does not work this way—for hunger makes eating attractive, not distastefully compulsive” (p. 127). Motivation normally depends on desire, not drive: “Philosophers since Aristotle have emphasized that desire is not a blind urge but rather represents its object as an ‘apparent good’ ...or under a ‘desirability characterization’ ...an attractive prospect that can elicit motivation to seek it—‘liking’ a representation gives rise to ‘wanting’ its object” (2013, p. 126; cf. Railton 2002; Berridge 2004; Aristotle 1999, §1113a15; Anscombe 1957, p. viii).

The point for now is that prospection nicely captures the forward-looking, “Garden of Forking Paths” character of deliberation and choice, since agents experience their options as a “branching array of evaluative prospects that fan out before them” (Seligman et al. 2013, p. 119).

Still, Gilbert and Wilson’s concerns remain: prospection seems limited to being episodic, since otherwise the simulated possibilities would overwhelm conscious mental activity. Moreover, prospection arguably lacks the “richness... of genuine perceptions” (2007, p. 1354). Why think that experiences of prospection are more common than Gilbert and Wilson allow, or that these experiences can compete with perceptual experience in being comparably motivational?

3.2 Prospection is richly experienced

Much of what Seligman et al. call “prospection” typically occurs outside conscious awareness and is unavailable to introspection, since it would be inefficient for agents to consciously keep track of all the simulations they generate. Yet prospection is often consciously experienced, and not just episodically. One reason it may be useful for prospection to be more continuously conscious is that it enables agents to engage in *shared* prospection. In other words, by communicating their simulations of future possibilities with one another, agents are able to draw on others’ cognition in deciding what to do. Moreover, many decisions are coordinated, and so agents must often prospect jointly, which is best achieved consciously (and often verbally). As a result, while it may be conceded that conscious prospection is less efficient than its unconscious cousin, more sustained (non-episodic) conscious prospection will enable agents to deliberate more effectively.

In the mechanism by which prospection becomes conscious, *affect* plays a central role. When prospection encounters “incommensurable dimensions and conflicting values and perspectives” (Seligman et al. 2013, p. 131), explicit comparison of these factors is facilitated by the “common metric” of affect:

Affect is the brain’s common currency for value, and conscious, subjective affect would permit the possible futures to be brought into the open for explicit comparison with each other. ...conscious subjective affect attached to prospectings would enable them to compete effectively with ongoing experience. (Seligman et al. 2013, p. 131)

Thus, whenever agents have conflicting thoughts about what to do, their simulated options feed into “an experientially rich and detailed workspace,” so that agents can “use their intelligence and imagination to best effect.” In such cases, “it can be best to act in *awareness* [emphasis added] of ...conflicting thoughts” (2013, p. 131). As a result, prospection becomes consciously experienced, and not just in sporadic episodes. Still, the experience of these simulations is not perceptual, since it does not even purport to furnish the agent with information about occurrent events inside or outside her body. As a result, we can grant the grain of truth in Gilbert and Wilson’s assertion that experiences of prospection fall short of “genuine perceptions” (2007, p. 1354). Yet we can also agree with the claim made by Seligman et al. that experiences of prospection occur frequently and function as competitors to

ongoing perceptual experience in being similarly motivational for the deliberating agent.

Prospection thereby explains how agents experience mere possibilities for choice. However, it fails to explain how agents might experience choice as indeterministic. Conscious prospection may feel free, since the prospecting agent's mind is—in some sense—freely exploring possibilities. Yet this implies nothing about indeterminism. I maintain that when we understand prospection in terms of causal modeling, we get the missing, seemingly indeterministic component.

3.3 Prospected choices are free variables

It is natural to think about the hypotheticals generated in prospection as carrying causal information about what would happen under variations in the values of variables—namely, alternative decisions the agent might make—in a causal model. Arguably, this is the only way in which we can adequately understand the information that is encoded in such hypotheticals (cf. Pearl 2001, p. 27). A causal model, in the present sense, is a representation that encodes hypothetical relationships between variables, where the variables represent causal relata. There is accumulating evidence that people's ordinary causal cognition is underpinned by this sort of modeling (Lagnado et al. 2013; Sloman 2005). Additionally, there is evidence from neuroimaging studies indicating that the same structures that are engaged by prospection are also used in the sort of counterfactual reasoning that is required for causal modeling (De Brigard et al. 2013).

On this way of thinking about causality and causal explanation, when we consider whether something is a cause, we ask, "What if things were different?" and by answering this question we identify factors whose manipulation would produce changes in the outcome being explained: If this (cause) variable were altered in these specific ways, then this other (effect) variable would be altered in these other ways. In the formal framework of causal modeling, the main restriction on what counts as a variable is that it must represent a particular event in such a way that it can be set to different values (different ways the event might be) by interventions (Woodward 2003, pp. 11–14).

An *intervention* is an "ideal" exogenous manipulation of a causal variable X in order to find out whether X causes Y. In short, we consider what would happen in the model by tweaking just X's value, and if a change reliably occurs in the value of Y, then X is judged to be a cause of Y. By contrast, endogenous changes occur because of the values already taken by other variables in the model. Interventions are "surgical," in that the usual causes of a variable, or of its taking a given value, are ignored or "screened off." In this procedure, we allow the variables whose antecedent causes we have screened off to vary freely across a range of values, where this range has a pragmatic restriction: none of the values should correspond to possibilities that we consider too remote (Woodward 2003, p. 90; cf. Hitchcock 2001, p. 287). On this approach, "X causes Y" (a type-level claim about "direct" causation) means that, for at least some state of the model, there is an intervention on X that would reliably change the value of Y.

Another idea that is central to the interventionist framework—along with that of an intervention—is *invariance*. There are two important components to invariance. First, for a generalization, *G*, relating two variables *X* and *Y* to be invariant, *G* must continue to describe how *Y* would change under a range of interventions on the value of *X*. Second, the generalization must also be invariant under a range of alterations in the background conditions, *C*.⁶ The invariance relations are stated as structural equations, which are “structural” or asymmetrical in the following sense: the value of the variable appearing on the left hand side of the “=” is determined by what appears on the right hand side:

$$Y = f(X, \dots, Z_n),$$

where the Z_n (if any) are just the other causal variables relevant to *Y* that are represented in the model. A function, *f*, for any structural equation is expressed as a mathematical operation on the variables. The equations comprising a model thereby encode (often multiple) counterfactuals of the following form:

If it were the case that the value of $X = x, \dots$, and $Z_n = z_n$, then it would be the case that the value of $Y = f(x, \dots, z_n)$, in background conditions *C*.

The background conditions are all the direct causes of *Y* that are not explicitly represented in the model. In this way, a causal model consists in a set of equations, which express the invariance relations obtaining between the variables, and which encode rich hypothetical or counterfactual causal information.

Token causation is defined in terms of a directed path, or “sequence,” of direct (type-level) causes. Relative to a causal model, “*X*’s taking the value *x* is an actual cause of *Y*’s taking the value *y*” means that there is at least one sequence of variables from *X* to *Y* (which may include only *X* and *Y*, or multiple intermediate variables as well) for which an intervention on *X* will change the value of *Y*, given that other direct causes of *Y* within the model and that are not part of this sequence are held fixed at their actual values.⁷

Philosophers have recently begun to use this sort of modeling to illuminate various questions about free will (e.g., Nahmias and Deery in preparation; Rose and Nichols 2013; Roskies 2012). In particular, Ismael (2013) claims that agents construct models like this when deliberating about what to do. In trying to decide among options for action, an agent’s cognition mentally carves off the event of her decision from its actual causal antecedents, and treats it as an exogenous variable in a causal model. By doing so, she is enabled to assess the “downstream” effects of this variable’s varying across a range of values, and this procedure yields causal information that is relevant to action-planning. These are the very hypotheticals that prospection generates in regulating the agent’s interaction with the environment.

Prospection in deliberative contexts plausibly treats the agent’s decision as an exogenous variable in a causal model, by screening off the decision’s causal

⁶ I will say more about why the background conditions are important in Sect. 4.

⁷ Cf. Woodward (2003, p. 42). This is what Yablo (2002, p. 138) calls “de facto” causal dependence of *Y* on *X*.

antecedents. Of course, most variables can be regarded as either exogenous or endogenous, depending on the context. An agent's decision might be treated as endogenous within a model of that decision developed by cognitive scientists, for instance. My claim is just that in first-personal deliberative contexts, an agent's prospected decision is naturally treated as exogenous—that is, as a variable whose prior causes have been screened off. For present purposes, this has a twofold significance. First, because such modeling treats the variable representing the decision as exogenous, it becomes easy for agents to experience (and to think of) their alternative options as indeterministic, as I explain in a moment. Second, in generating simulated possibilities, agents not only pay less attention to a decision's causal antecedents and more attention to what Seligman et al. call a “branching array of evaluative prospects that fan out before them” (2013, p. 119). Mentally, they *ignore* the antecedent causes, and so are liable to experience the decision as lacking deterministic causes.

3.4 Indeterminist experience and belief

While Seligman et al. (2013) weigh in briefly on the topic of free will, they do not address the sort of evidence that I considered earlier, which indicates that people tend to experience possessing, and to believe that they possess, an indeterminist freedom. Even so, I maintain that prospection—together with an interventionist account of how the hypotheticals generated by prospection are to be modeled—explains why people tend to believe in indeterminist freedom, and for two reasons.

First, whenever agents generate simulated possibilities for action while deliberating about what to do, the variable representing their decision is treated by prospection as free, meaning that it is permitted to vary over a range of values. Yet, were the deliberating agent to consider the same decision while assuming determinism, she would instead treat it as having antecedent sufficient causes, and thus as a variable the values of which are constrained by the wider model of the deterministic system. Here, the variable is permitted to take just a single value.⁸ This creates an apparent psychological conflict between treating one and the same variable as free and constrained, which makes the experience seem indeterminist. When the agent tries to hold in mind both models of her decision at the same time—say, in a forced choice experiment in which she is asked whether her experience is consistent with determinism (cf. Deery et al. 2013)—each model will be felt to be inconsistent with the other.

It is important to be clear about what I am claiming is going on in such cases. Broadly speaking, there are two options to consider regarding what is going on when an agent judges her experience as incompatibilist. First, it might be that the experience does not actually require indeterminism to be accurate, but somehow the agent mistakenly judges that it does (cf. Horgan 2011, 2012). Second, it might be

⁸ Of course, this value may be unknown—or even unknowable—prior to the agent's actually deciding. This is suggestive of epistemic possibility, which is entirely compatible with determinism. I have more to say on this issue shortly.

that the experience *does* require indeterminism to be accurate,⁹ yet—due to other considerations—the agent’s forming a belief in indeterminism is not justified by the experience.

My suggestion lies between these two options. I claim that when an agent reports having an indeterminist experience, and forms a belief in indeterminism on this basis, her belief is not justified since the openness of the experience is ultimately to be cashed out in terms of epistemic possibility—though not (as I will explain) in terms of *ordinary* epistemic possibility. To maintain that it is epistemically open whether I will choose either of two options, such as going left or right, is just to maintain that my choosing either option is consistent with what I know (Pereboom, 2008, pp. 292–296; cf. Kapitan 1986; Nelkin 2004). Famously, Smart (1961) claimed that this is how we often interpret counterfactuals outside the sphere of action. When I say, “the plate fell, and it could have broken,” Smart maintains, I am not making a claim about determinism. All I am saying is that before the plate completed its fall, for all I knew it would break (1961, p. 298). Similarly, if we say that Lee Harvey Oswald could have done otherwise than shoot President Kennedy, all we are saying is that before Oswald pulled the trigger, for all we knew he would not pull it. In making claims about epistemic possibility, there is no conflict with determinism. As a result, explaining experiences of freedom in terms of epistemic possibility not only fails to support a belief in indeterminism, but also fails to support the view that the *experience* itself is indeterminist. For this reason, I appeal to a “richer” variety of epistemic possibility, which I explain in terms of causal modeling.

The rich epistemic possibility to which I appeal is liable to be judged as inconsistent with determinism because in screening off her decision from its antecedent causes, an agent’s modeling of that decision in prospect screens off a large part of what she actually knows, or might reasonably be expected to bring to mind in other contexts—for example, that her decisions have prior causes. So the epistemic possibilities available to the agent when she is deliberating have a restriction on the “for all I know” that yields ordinary epistemic possibility. This permits the agent to experience her available options as more robustly open than they would otherwise seem to be.

In other words, while the agent is—in a sense—misinterpreting her experience of openness when she judges it as indeterminist, matters are not as straightforward as some philosophers who maintain that there is nothing in the experience that is suggestive of indeterminism make them out to be.¹⁰ On my view, the suggestion of indeterministic openness is *built right into the experience*, as a result of the rich epistemic possibility at work in the modeling that generates the experience. Still, as I explain in Sect. 4, this suggestiveness cannot be used to justify a belief in indeterminism, since the openness is still epistemic in nature. Whereas a model of one’s prospected decision treats the variable representing that decision as varying over a range of values, the permissible values for this variable will be constrained to

⁹ See Deery (2014) for the view that even experiences with indeterminist content might be accurate, assuming determinism.

¹⁰ This, as I understand it, is Horgan’s view (e.g., 2011, 2012).

just one by the wider model of determinism. While these models might be *felt* as inconsistent, actually they are not. As Ismael (2013) puts it,

Our actions appear in multiple models, sometimes as exogenous, sometimes as endogenous. Whenever we model a subsystem of the world... we can always widen our view to attain a more encompassing perspective, and wide-scope models will have a set of possibilities that is typically constrained relative to the narrow-scope model... But there is no more conflict between these models than there is between the view of a building from close-up and the view from a very great distance. (2013, p. 230)

Indeed, the foregoing considerations indicate a further reason why people's experience may seem indeterministic. If prospection treats choice as a free variable, then agents not only pay less attention to the causal antecedents of their decisions, but mentally they ignore such causes. In other words, agents experience their prospected decisions as *not* determined by anything prior to them. In modeling a given decision as free, prospection screens it off from its prior causes, allowing it to vary freely across a range of values. In this way, the decision is modeled as not having antecedent sufficient causes, and—since this modeling *generates* the experience—the experience also represents the decision as lacking such causes. A decision that is experienced as not having antecedent sufficient causes is, *a fortiori*, experienced as not having antecedent deterministic causes.

If either of these explanations is on the right track, then people's belief in indeterminist free will plausibly has its source in their deliberative experience. Whereas causal modeling by itself fails to indicate a mechanism by which people might come to believe in indeterminist freedom, prospection yields the missing mechanism. Affect brings the simulated possibilities into conscious awareness, in a way that competes with ongoing perceptual experience. Since the modeling of these possibilities enables people to experience them as indeterministic, we also have a mechanism by which people can experience decisions as indeterministic. People tend to believe that the future is open in a manner that would require indeterminism for the belief to be accurate, since that is what is suggested by their experience. People may also tend to believe that their decision is not sufficiently caused by anything prior to it, and thus (implicitly) that it lacks deterministic causes, since that, too, is what is suggested in experience. This belief might be defeated once people learn that their decision *does* have prior sufficient causes—perhaps even deterministic ones. Nevertheless, the initial tendency may be to believe that it does not. And the experience itself will presumably remain unaltered.

Thus, contra Nichols' claim that experience is too "anemic" to be a source of the belief in indeterminism, such a belief may well come from people's ordinary experience of deliberating and choosing.

4 Putting the theory to work

The account I have outlined explains why people tend to believe in indeterminist freedom. Yet it also does more than that, since it explains why decisions are treated

as special in this regard, and why people do not as easily think of other contingent events as requiring indeterminism.

As noted already, there is excellent evidence that people's causal cognition employs causal modeling. Yet people certainly do not have indeterminist beliefs about *all* causal relations. Typically, people do not even have indeterminist beliefs about events that they consider contingent, such as whether a leaf floating down a stream will go left or right, or whether a storm will hit a city, or whether a wineglass will break when a gust of wind knocks it from the table onto a hardwood floor. Additionally, people's modeling of mental processes in general is not indeterministic. We might therefore ask why decisions are special in being experienced and judged as indeterministic.

Recent experimental work conducted by Lombrozo (2010) suggests that the way in which people's causal cognition differs across such cases is according to whether the events in question are described teleologically (in terms of goals) or instead mechanistically (for instance, in terms of physical connection or the transfer of a quantity of energy). When participants are asked to explain why an event construed teleologically occurs, typically they are sensitive to causal difference-making of the sort that is captured by causal modeling. Yet when events are construed mechanistically, participants are mostly sensitive to physical connection. This suggests that when people consider whether a leaf floating down a stream will go left or right, mechanistic thinking takes precedence. People judge that the leaf will go a particular way, given the physical forces impacting on it. By contrast, when people consider decisions, their thinking works differently. Imagine a skier who faces two options: go left or right around a rock on a mountainside. Here, people tend to think less in terms of physical forces determining the outcome, and more in terms of how the outcome depends on the skier's goals at the moment of her decision. In other words, people are less likely to think about the floating leaf, and more likely to think about the skier's decision, in terms of the rich counterfactual dependence captured by causal modeling. James (1890) nicely illustrates this difference between intentional agents and non-intentional objects:

Romeo wants Juliet as the filings want the magnet; and if no obstacles intervene he moves towards her by as straight a line as they. But Romeo and Juliet, if a wall be built between them, do not remain idiotically pressing their faces against its opposite sides like the magnet and the filings [when a card is placed between them]. Romeo soon finds a circuitous way, by scaling the wall or otherwise, of touching Juliet's lips directly. With the filings the path is fixed; whether it reaches the end depends on accidents. With the lover it is the end which is fixed, the path may be modified indefinitely. (1890, p. 20)

According to Lombrozo (2010, pp. 309–310), actions like Romeo's exhibit *equipfinality*—the same outcome would be achieved by different means. The behavior of the iron filings, by contrast, exhibits *multifinality*—different means would result in different outcomes. Consideration of agents' intentional actions emphasizes counterfactual dependence relationships, whereas consideration of the behaviors of non-intentional objects, like iron filings or floating leaves, does not (cf. Lagnado and Channon 2008; Malle 2004; McClure et al. 2007). My account of

the source of indeterminist beliefs relies on the idea that people think of contingent choices as equifinal, yet mechanistically described events as multifinal.

Nevertheless, people sometimes treat multifinal events as contingent. After all, even when the leaf goes left, it still makes sense to ask whether it could have gone right. And presumably, all else being equal, people typically judge that it might have. But even if cognition utilizes causal modeling in this sort of case, there is an important reason why people are less likely to form indeterminist beliefs about such an event, as I outline in a moment. Similarly, events involving non-intentional objects might exhibit equifinality to a degree, yet not be modeled as indeterministic—for instance, a plant’s growing to the left instead of the right in order to avoid an obstacle that is blocking its sunlight.¹¹ The plant’s going left is contingent, yet is at least weakly equifinal (unlike the leaf in the stream). Even so, the plant’s ability to avoid obstacles that block its sunlight is nowhere near as strong as Romeo’s ability to overcome obstacles that prevent his reaching Juliet. The difference lies in the *strength of invariance* obtaining between each of these causes and its corresponding effect. For present purposes, let us say that an invariance relation, R_1 , between two variables X and Y is stronger than another such relation, R_2 , between Y and its prior causal variables (including X) if:

- (1) holding fixed the relevant background conditions, C , R_1 predicts the value of Y under a wider range of interventions on X than R_2 does; and
- (2) R_1 predicts the value of Y across a wider range of changes to the values of C than R_2 does.¹²

What clause (2) says is that the counterfactuals encoded in the R_1 equation—which expresses the invariance relation obtaining between X and Y —predict the value of Y under a wider range of counterfactual situations, or changes in background conditions, than those encoded in the R_2 equation. So, an alarm clock that is able to switch to a battery backup mechanism during a power cut will have a stronger invariance relation between its being set to go off and its going off than a clock without a battery. This is because the clock with the battery will go off across a wider range of changes in background conditions (such as power cuts) than a clock without a battery. Similarly, the invariance relation between Romeo’s deliberation and his reaching Juliet is of a stronger *type* than the one between the plant’s sunlight being blocked and its reaching the light by some other means. Certain obstacles, or changes in background conditions, may stop the plant, but not Romeo. As Lombrozo puts it, “Goal-directed human behavior is a gold standard for equifinality” (2010, p. 310).

My main claim is that equifinal human choices are more likely than multifinal contingent events to be judged as indeterministic. Yet I also claim that equifinal choices are more likely to be judged as indeterministic than other equifinal contingent events that have weaker invariance relations to their effects. Finally, an

¹¹ I thank an anonymous reviewer for this example.

¹² These conditions suffice for present purposes, but are simplified. See Woodward (2003, Chap. 6) for an extended discussion.

agent's own deliberations and decisions, due to the way in which prospecting enables the agent to experience her decision from the first-person perspective, are the most likely events of all to be modeled as indeterministic, for the reasons I have already outlined.¹³

Finally, one might think that while my account explains what Nahmias (2006) calls “close call” decisions—that is, “situations in which, after deliberating, we still do not know what we really want to do” (2006, p. 627)—it does not explain another candidate for paradigmatically free choice, namely the decisions of an agent who, “after deliberating, ...feels confident about what she really wants..., chooses it, and then effectively acts on her choice” (2006, p. 630). When an agent is confident, she does not experience her future as open—let alone indeterministically open—since that would undermine her confidence. Since confident decisions are, Nahmias thinks, more paradigmatically free than close-call decisions, my account fails to explain the most paradigmatically free of decisions.

However, while close-call decisions are plausibly experienced as more radically indeterministic than confident decisions, confident decisions might still be experienced as indeterministic. Consider the following case: I am mugged at gunpoint, and prudently I decide to hand over my wallet. Despite my confidence in this decision, I might still experience myself as free to refuse to hand over the wallet (cf. Horgan 2007, p. 22, note 24). Presumably, this will be so even if I judge it irrational to refuse. Unless we permit the variable representing my decision to range over more than one value, I cannot model (or experience) the episode as a free decision at all. And once I do model it as free, my prospecting is likely to bring the decision into experience as seemingly indeterministic. Models on which the variable representing my decision is permitted to take just one value are not the sorts of models in which prospecting deals. Consequently, the account I have developed can allow that some decisions are more confident than others, while also explaining the common feeling—as in the mugging case—that we are free to perform actions we judge as irrational.

All of this accords closely with libertarian descriptions of the experience of free will. However, even if people really do have indeterministic experiences of freedom, it does not follow that these experiences are accurate, or that a belief in indeterminism formed on their basis is justified.

First, if the etiology I have sketched for indeterministic experiences is right, all that is actually going on is that people are considering what outcomes they can cause, depending on the decisions they make, and the modeling involved requires that they let the event of their decision range over more than one value. People simulate the possible effects of their deciding in different ways, and this process is consistent with determinism. Even for a deterministic system in which all events are represented by variables that may take just a single value (once the initial conditions and global laws are fixed), it still makes sense to ask what *would* happen in a sub-

¹³ Even complex models of other agents' decisions are comparatively less likely to result in indeterministic experiences or beliefs than one's own prospected decisions are, since it is more difficult to screen off other agents' decisions from their prior causes. Plausibly, people think more *deterministically* when they consider other agents' decisions (Hume 1955/1743, pp. 83, 88; Nichols 2004).

model of that system were we to permit a variable to take another value. That is simply to employ counterfactual reasoning, and to ask how events would unfold if something antecedent to that event were different.

Second, even if people's introspection seems to reveal that one's decision lacks deterministic causes, it does not follow that it actually lacks such causes, since the causes may not be introspectible. At bottom, this is a logical point. Yet the suggestion that people are ignorant of the deterministic causes of their decisions has a long history (cf. Spinoza 1887/1677, pp. 108, 134; D'Holbach 1970/1770, Part I, Chap. 11). More recently, there is evidence indicating that the best predictors of many decisions are not introspected by the agents themselves, and in some cases may not even be introspectible (e.g., Mathews and Cannon 1975; Darley and Batson 1973; Isen and Levin 1972; Libet 1999; Wegner 2003). If the world is deterministic, these influences function as deterministic causes of decisions, in which case deterministic causes of decisions are non-introspectible.

Third, the hypotheticals generated in prospecting cannot support the view that the future is in fact indeterministically open. This is due to the epistemic nature of such hypotheticals, as we have seen.

5 Rival accounts

Other explanations of the folk belief in indeterminist freedom have been proposed, yet have significant shortcomings. As I noted earlier, Nichols (2012) considers it implausible to think that indeterminist beliefs have their source in experience, since indeterminism is, Nichols thinks, too sophisticated a notion to grasp on the basis of raw experience—that is, experience unaided by beliefs. Nichols grants that experience might be indeterminist, but if so, indeterminist experience is caused by indeterminist belief.

Nichols' suggestion is that background beliefs might alter experiences of freedom via cognitive penetration. Roughly, cognitive penetration occurs when the character of one's experience is altered by one's cognitive states—for instance, by one's background beliefs. If two participants report different visual experiences while looking at the same patch of color under the same conditions, the thesis of cognitive penetration explains this variation by appeal to differences in the participants' cognitive states. While controversial, there is evidence that cognitive penetration occurs in visual experience (e.g., Levin and Banaji 2006; Delk and Fillenbaum 1965; cf. McCauley and Henrich 2006). If visual experience is cognitively penetrable, then in the absence of any reason to think that cognitive penetration could not also occur for experiences of freedom, Nichols thinks it plausible that background beliefs might influence these experiences too. As a result, even if experiences of freedom are not initially indeterminist in character, that does not mean that they remain that way: background beliefs might alter the experiences, making them indeterminist.

One might think that cognitive penetration is implausible for experiences of freedom. Yet even if such experiences are penetrable, that would not explain indeterminist belief, since it could do so only by appeal to experience, which is itself

shaped by belief. As Nichols puts it, “If experience is supposed to provide a noncircular explanation for our belief in indeterminism (or our theoretical resistance to determinism), then it has to be in virtue of experience that is not guided by an indeterminist belief” (2012, p. 294).

To avoid this problem, one might reconsider the possibility that Nichols rejects—namely, that indeterminist beliefs have their source in experience.¹⁴ Holton (2006) maintains, for instance, that “Our experience tells us that our choice is not determined by our beliefs and desires, or by any other psychological states ...to which we have access. Those could be the same, and yet we could choose differently” (p. 15). As a result of such experience, people tend to believe indeterminism.

Horgan (2011, 2012) agrees with Holton that people tend to form indeterminist beliefs on the basis of their reported experience, yet he denies that these experiences are indeterminist. Horgan maintains that people might *think* they can tell by introspection whether their experience is indeterminist. Yet this reflects an introspective error. As Horgan puts it, it is one thing to know (A) by introspection:

(A) My experience does *not* present my choice as determined by my prior states.

Yet it is another thing to know (B) by introspection:

(B) My experience presents my choice as *not* determined by my prior states.

According to Horgan, it is plausible that agents can discern whether (A) is true by introspection, but not whether (B) is true. When one asserts (B), and as a result judges one’s experience as indeterministic, either one is mistakenly inferring (B) from (A), or one is simply conflating (A) and (B). Either way, a resulting belief in indeterminism will be fallaciously formed.

Nichols (2012) thinks that even Horgan’s more sophisticated etiology for the belief in indeterminism cannot be right, since it fails as an explanation of the belief. We are not introspectively aware of the causes of mental states like headaches, Nichols observes, but we do not infer indeterminism from that:

It would be a kind of scope fallacy to move from “I don’t experience my actions as determined” to “I experience my actions as not determined.” Now, people surely do commit fallacies. But notice we don’t seem to commit the scope fallacy when it comes to headaches. That is, the phenomenology of headaches doesn’t present us with a set of deterministic headache-causes, but we don’t leap to indeterminist conclusions there. (2012, p. 296)

Nichols proposes an alternative etiology for indeterminist beliefs. First, he outlines how scientists are able to judge whether a system is indeterministic. Researchers control for inputs to the system, and if it has different outputs given the same inputs, the researchers are warranted in concluding that the system is indeterministic (2012,

¹⁴ This move is further motivated by the fact that there is no positive evidence that cognitive penetration occurs in agentive experience.

p. 297). Next, Nichols presents an argument for indeterminism adapted from William of Ockham (2012, pp. 298–299):

- (1) The factors that are introspectively accessible to me do not determine my choice.
- (2) I have introspective access to all the (proximal) factors that influence my choice.
- (3) Therefore, my choice is not determined.

In this argument, (2) is the crucial premise, and one might deny that people believe it. However, Nichols notes that we should not interpret (2) as a stronger claim than it is. After all, (2) does not claim that people assume they have access to *everything* in their minds. It is a more limited claim: “All that is required is a kind of default (but defeasible) assumption that the causal influences *on decisions* [emphasis added] are introspectively available. And there is evidence that people do have such a default assumption” (2012, p. 299). So people tend to believe (2). Further, decisions are special in this regard—people do not make the same assumption when it comes to whether all the proximal factors influencing their urges (for example) are introspectively accessible. Presumably, the same goes for headaches.

Why do people make the default assumption for decisions, yet not for urges or headaches? Nichols thinks it is because of a certain bias: “Our attention is drawn to factors that are present, rather than to the possibility that there are hidden factors” (2012, p. 301). Normally, people take themselves to understand how mechanisms with discrete, accessible causal parts work—for instance, locks or zippers. (These causal parts are “present.”) Yet when asked to explain how locks or zippers work, people typically find themselves at a loss. As a result, they downgrade their level of presumed understanding. By contrast, people do not take themselves to have a good understanding of how complex objects like flash drives work, since they lack easy access to the causal parts of such devices. (These causal parts are “hidden.”) Nichols thinks that when people introspect on decisions, they find that they have access to discrete mental states (beliefs, thoughts, desires) that causally influence their decisions. So they take decisions to be like locks or zippers. When people are presented with evidence from psychology detailing the actual, unconscious causes of decisions, they downgrade their presumed level of understanding, and can be brought to reject (2). People’s default setting is to think of decisions as like zippers, but of headaches as like flash drives. Yet just as people can be brought to downgrade their presumed understanding of zippers, so too they can be brought to downgrade their understanding of decisions.

This is meant to explain why people tend to believe premise (2)—the claim that people think they have introspective access to all the proximal factors that influence their decisions. Of course, it also shows that belief in indeterminism is not warranted, at least if it results from anything like the inference from premises (1) and (2) to (3). This is because people’s default assumption that they have introspective access to all the factors that influence their decisions is mistaken.

Nichols view is attractive and may indeed be part of the story about why people tend to believe indeterminism. Yet a shortcoming of Nichols’ proposal is its backward-looking orientation: it focuses solely on individuals’ access to the causes

of their decisions. This is true of Horgan's and Holton's accounts as well. At bottom, however, deliberation and choice are forward-looking phenomena. When agents look to the future when deliberating about what to do, two things are typically salient for them, which Holton, Horgan, and Nichols all fail adequately to address: the experience of having alternative possibilities for action, and the experience of being free to decide between such alternatives. The world seems, in other words, to be a Garden of Forking Paths. Indeed, these are some of the features of experiences of freedom that libertarians most often cite as indeterministic.

Against Nichols' proposed explanation for the belief in indeterminism, I maintain that this belief instead has its source in people's deliberative experiences of navigating into the future, rather than from their introspection on the causes of their decisions. People's experience of prospection, because of the way in which the hypotheticals generated in prospection are produced, results in a belief in indeterminist freedom that derives directly from experience. Moreover, these indeterminist beliefs do not result from a mistaken judgment about experience, as Horgan claims. Finally, although indeterminist beliefs partly derive, as Holton rightly maintains, from a feeling that one's choice is not determined, this is due in the first instance to the forward-looking nature of deliberative experience, rather than its backward-looking character. Nevertheless, while it may be natural for people to arrive at indeterminist beliefs on the basis of their experience, these beliefs are not justified.

A more direct rival to my view claims that indeterminist beliefs have their source in perceptual experience. A number of researchers have developed forward-looking, quasi-perceptual models of agentive experience (e.g., Bayne 2008, 2011; Synofzik et al. 2008; Bayne and Pacherie 2007; Pacherie 2007; Blakemore and Frith 2003). These authors maintain that experience enables agents—minimally—to distinguish self-generated actions from involuntary bodily movements and the externally caused movements of objects. Bayne (2011) develops a prominent view of this sort.

On Bayne's view, agentive experiences are produced by a dedicated perceptual system, which includes forward models of action control. These models receive a copy of the agent's motor commands, which are used to predict the sensory consequences of the agent's bodily movements. A copy of the motor commands is also sent to a "comparator," so-called because it compares the predicted sensory consequences of bodily movement with sensory feedback.¹⁵ When the comparator finds a match between prediction and feedback, the movement is identified as self-generated. When there is no match (or a weak match), the movement is identified as externally caused. Thus, the comparator model explains how agents are able (experientially) to distinguish self-generated actions from involuntary bodily movements and the externally caused movements of objects.

¹⁵ Bayne locates the comparator in cognitive architecture *between* the standard perceptual systems and the motor system, since it takes both perceptual and motor representations as inputs. The states generated by the comparators have the functional role of perception. Like other perceptions, the function of this "sense of agency" is to generate representations of some domain and make these available to the agent's cognitive systems in an experiential format (2011, pp. 358–359).

A central shortcoming of Bayne's view, however, is that it cannot explain experiences of freedom, let alone how a belief in freedom might derive from such experience. Nevertheless, Bayne thinks that we experience freedom: "[T]here is clearly an intuitive sense in which we can—and often do—experience ourselves as acting freely," although "[t]he difficulties that confront us in attempting to articulate what it is like to experience oneself as a free agent are ...particularly imposing" (2008, pp. 195–196). Indeed, Bayne's model falls especially short in explaining experiences of *indeterminist* freedom, since he thinks it unclear whether such experiences are possible:

[P]erhaps experiential systems are incapable of inserting a negation quite where it needs to be inserted in order to represent libertarian freedom. In order to represent an action as free in a libertarian sense one must not only represent it as undetermined by one's prior psychological properties but also as undetermined by one's physical properties—or indeed any physical properties. And it is not obvious that experiential systems have *that* kind of representational power. (2008, p. 196)

This echoes Nichols' and Horgan's suggestion that even if people do *not* experience their choices as determined, it is unlikely that they experience them as *not* determined. Of course, this slips into the backward-looking perspective emphasized by Nichols and Horgan,¹⁶ which I claim fails to capture the centrally forward-looking, or Garden-of-Forking-Paths character of deliberative experience.

In any case, Bayne's view fails to explain indeterminist experiences of freedom, or how a belief in indeterminism might have its source in experience. This, I suggest, is because his model focuses on perceptual experience, and it is difficult to see how perception could have as content that one is free to do otherwise, even in the sense that one could do otherwise *if* one wanted—a sense that is consistent with determinism.¹⁷ To borrow Nichols' phrase, perceptual experience is too "anemic." To have an experience of freedom, an agent must compare two or more distinct representations—the alternative possibilities themselves—in the mind, and that is not a perceptual operation.

Although Bayne is right to focus on the forward-looking aspect of agency, his account does not explain the sort of experience that might plausibly generate a belief in indeterminism. Yet contra Nichols, I maintain that experience is the right place to seek an explanation of people's tendency to believe in indeterminism. The relevant kind of experience, however, is quite different from the sort on which Bayne focuses. The experience driving belief in indeterminism results from

¹⁶ Interestingly, Bayne's backward-looking suggestion for why it may be impossible to experience indeterminist freedom is at odds with his own forward-looking framework, since—like Nichols' and Horgan's views—it focuses on the causes of decisions.

¹⁷ Consequently, it is unclear how any development of Bayne's view could account for the sort of freedom at issue in free-will debates. No one thinks *general* abilities, for instance, are incompatibilist, or that one's having *kinds* of actions one can perform is inconsistent with determinism, and the relevant experiences presumably might be accurate, even assuming determinism. Thus, any development of Bayne's view along these lines seems bound to fall short of accounting for any variety of freedom that might be a matter of contention between compatibilists and incompatibilists.

prospection, which is the mental simulation of future possibilities for the purpose of guiding action. These simulations function as competitors to perceptual experience, in such a way that they are genuinely experiential. Finally, the way in which prospection causally models possibilities for action makes it easy for deliberating agents to experience—and to believe—that their freedom is indeterministic.

6 Conclusion

The view I have developed accomplishes two things. First, it explains why people tend to believe that their freedom to decide among alternatives is indeterministic: they do so because of how they experience these decisions, due to the phenomenon of prospection and the causal modeling that drives it. Second, my view explains how indeterminist experiences fail to justify a belief in libertarianism—the view that freedom requires indeterminism and we are free. This is because the experience of prospection is ultimately of a sort of epistemic openness. As a result, my view provides a deflationary explanation of a traditional motivation for libertarianism, while nevertheless conceding that libertarians are in fact right about something: namely, that it is tantalizingly easy for deliberating agents to experience—and thus to believe in—indeterminist free will.

Acknowledgments Thanks to Paul Russell, Taylor Davis, Eddy Nahmias, Jenann Ismael, Terry Horgan, Shaun Nichols, Bryan Chambliss, Will Leonard, and an anonymous referee at this journal for helpful feedback and comments.

References

- Anscombe, G. E. M. (1957). *Intention*. Oxford: Blackwell.
- Aristotle. (1999). *Nicomachean ethics* (T. Irwin, Trans.). Indianapolis: Hackett.
- Bayne, T. (2008). The phenomenology of agency. *Philosophy Compass*, 3(1), 182–202.
- Bayne, T. (2011). The sense of agency. In F. Macpherson (Ed.), *The senses* (pp. 355–374). Oxford: Oxford University Press.
- Bayne, T., & Pacherie, E. (2007). Narrators and comparators: The architecture of agential self awareness. *Synthese*, 159(3), 475–491.
- Berridge, K. (2004). Motivation concepts in behavioral neuroscience. *Physiology & Behavior*, 81, 179–209.
- Blakemore, S.-J., & Frith, C. (2003). Self-awareness and action. *Current Opinion in Neurobiology*, 13(2), 219–224.
- Buckner, R., & Carroll, D. (2007). Self-projection and the brain. *Trends in Cognitive Science*, 11, 49–57.
- Campbell, C. A. (1951). Is 'freewill' a pseudo-problem? *Mind*, 60(240), 441–465.
- Darley, J. M., & Batson, C. D. (1973). From Jerusalem to Jericho: A study of situational and dispositional variables in helping behavior. *Journal of Personality and Social Psychology*, 27, 100–108.
- De Brigard, F., Addis, D. R., Ford, J. H., Schacter, D. L., & Giovanello, K. S. (2013). Remembering what could have happened: Neural correlates of episodic counterfactual thinking. *Neuropsychologia*, 51(12), 2401–2414.
- Deery, O. (2014). The fall from Eden: Why libertarianism isn't justified by experience. *Australasian Journal of Philosophy*. doi:10.1080/00048402.2014.968596
- Deery, O., Bedke, M., & Nichols, S. (2013). Phenomenal abilities: Incompatibilism and the experience of agency. In D. Shoemaker (Ed.), *Oxford studies in agency and responsibility* (pp. 126–150). New York: Oxford University Press.

- Deery, O., Davis, T., & Carey, J. (2014). The free-will intuitions scale and the question of natural compatibilism. *Philosophical Psychology*. doi:10.1080/09515089.2014.893868
- Delk, J. L., & Fillenbaum, S. (1965). Differences in perceived color as a function of characteristic color. *The American Journal of Psychology*, 78(2), 290–293.
- D'Holbach, B. (1970/1770). *The system of nature* (H. D. Robinson, Trans.). New York: B. Franklin.
- Fara, M. (2008). Masked abilities and compatibilism. *Mind*, 117(468), 843–865.
- Feltz, A., & Cokely, E. (2009). Do judgments about freedom and responsibility depend on who you are? Personality differences in intuitions about compatibilism and incompatibilism. *Consciousness and Cognition*, 18(1), 342–350.
- Feltz, A., Perez, A., & Harris, M. (2012). Free will, causes, and decisions: Individual differences in written reports. *Journal of Consciousness Studies*, 19(9–10), 166–189.
- Fischer, J. M. (1994). *The metaphysics of free will*. Oxford: Blackwell.
- Gilbert, D., & Wilson, T. (2007). Propection: Experiencing the future. *Science*, 351, 1351–1354.
- Hitchcock, C. R. (2001). The intransitivity of causation revealed in equations and graphs. *Journal of Philosophy*, 98, 273–299.
- Holton, R. (2006). The act of choice. *Philosophers' Imprint*, 6(3), 1–15.
- Horgan, T. (2007). Agentive phenomenal intentionality and the limits of introspection. *Psyche*, 13, 1–29.
- Horgan, T. (2011). The phenomenology of agency and freedom: Lessons from introspection and lessons from its limits. *Humana Mente*, 15, 77–97.
- Horgan, T. (2012). Introspection about phenomenal consciousness: Running the gamut from infallibility to impotence. In D. Smithies & D. Stoljar (Eds.), *Introspection and consciousness* (pp. 403–422). Oxford: Oxford University Press.
- Hume, D. (1955/1743). *An enquiry concerning human understanding* (L. Selby-Bigge, Ed.). Oxford: Clarendon Press.
- Isen, A. M., & Levin, P. F. (1972). Effect of feeling good on helping: cookies and kindness. *Journal of Personality and Social Psychology*, 21, 384–388.
- Ismael, J. (2013). Causation, free will, and naturalism. In H. Kincaid, J. Ladyman, & D. Ross (Eds.), *Scientific metaphysics* (pp. 208–235). New York: Oxford University Press.
- James, W. (1890). *The principles of psychology*. Cambridge, MA: Harvard University Press.
- Kapitan, T. (1986). Deliberation and the presumption of open alternatives. *The Philosophical Quarterly*, 36(143), 230–251.
- Knobe, J. (2014). Free will and the scientific vision. In E. Machery & E. O'Neill (Eds.), *Current controversies in experimental philosophy* (pp. 69–85). New York: Routledge.
- Lagnado, D. A., & Channon, S. (2008). Judgments of cause and blame: The influence of intentionality and foreseeability. *Cognition*, 108, 754–770.
- Lagnado, D., Gerstenberg, T., & Zultan, R. (2013). Causal responsibility and counterfactuals. *Cognitive Science*, 37, 1036–1073.
- Levin, D. T., & Banaji, M. R. (2006). Distortions in the perceived lightness of faces: The role of race categories. *Journal of Experimental Psychology: General*, 135(4), 501–512.
- Libet, B. (1999). Do we have free will? *Journal of Consciousness Studies*, 6, 47–57.
- List, C. (2014). Free will, determinism, and the possibility of doing otherwise. *Noûs*, 48(1), 156–178.
- Lombrozo, T. (2010). Causal-explanatory pluralism: How intentions, functions, and mechanisms influence causal ascriptions. *Cognitive Psychology*, 61(4), 303–332.
- Malle, B. F. (2004). *How the mind explains behavior: Folk explanations, meaning, and social interaction*. Cambridge, MA: MIT Press.
- Mathews, K. E., & Cannon, L. K. (1975). Environmental noise level as a determinant of helping behavior. *Journal of Personality and Social Psychology*, 32, 571–577.
- May, J. (2014). On the very concept of free will. *Synthese*. doi:10.1007/s11229-014-0426-1.
- McCauley, R. N., & Henrich, J. (2006). Susceptibility to the Müller–Lyer illusion, theory neutral observation, and the diachronic cognitive penetrability of the visual input system. *Philosophical Psychology*, 19(1), 79–101.
- McClure, J., Hilton, D. J., & Sutton, R. M. (2007). Judgments of voluntary and physical causes in causal chains: Probabilistic and social functionalist criteria for attributions. *European Journal of Social Psychology*, 37, 879–901.
- Murray, D., & Nahmias, E. (2014). Explaining away incompatibilist intuitions. *Philosophy and Phenomenological Research*, 88(2), 434–467.
- Nahmias, E. (2006). Close calls and the confident agent: Free will, deliberation, and alternative possibilities. *Philosophical Studies*, 131(3), 627–667.

- Nahmias, E., Coates, J., & Kvaran, T. (2007). Free will, moral responsibility, and mechanism: Experiments on folk intuitions. *Midwest Studies in Philosophy*, 31(1), 214–242.
- Nahmias, E., & Deery, O. (In preparation). Defeating manipulation arguments: Interventionist causation and compatibilist sourcehood.
- Nahmias, E., Morris, S. G., Nadelhoffer, T., & Turner, J. (2004). The phenomenology of free will. *Journal of Consciousness Studies*, 11(7–8), 162–179.
- Nahmias, E., & Murray, D. (2011). Experimental philosophy on free will: An error theory for incompatibilist intuitions. In J. Aguilar, A. Buckareff, & K. Frankish (Eds.), *New waves in philosophy of action* (pp. 189–215). London: Palgrave-Macmillan.
- Nelkin, D. (2004). Deliberative alternatives. *Philosophical Topics*, 32, 215–240.
- Nichols, S. (2004). The folk psychology of free will: Fits and starts. *Mind and Language*, 19(5), 473–502.
- Nichols, S. (2012). The indeterminist intuition: Source and status. *The Monist*, 95(2), 290–307.
- Nichols, S., & Knobe, J. (2007). Moral responsibility and determinism: The cognitive science of folk intuitions. *Noûs*, 41(4), 663–685.
- O'Connor, T. (1995). Agent causation. In T. O'Connor (Ed.), *Agents, causes, and events: Essays on indeterminism and free will* (pp. 173–200). New York: Oxford University Press.
- Pacherie, E. (2007). The sense of control and the sense of agency. *Psyche*, 13(1), 1–30.
- Pearl, J. (2001). Bayesianism and causality, or, why I am only a half-Bayesian. In D. Cornfield & J. Williamson (Eds.), *Foundations of Bayesianism* (pp. 19–36). Dordrecht: Kluwer.
- Pereboom, D. (2008). A compatibilist account of the epistemic conditions on rational deliberation. *The Journal of Ethics*, 12, 287–306.
- Railton, P. (2002). Kant meets Aristotle where reason meets appetite. In C. U. Moulines & K.-G. Niebergall (Eds.), *Argument Und Analyse*. Paderborn: Mentis.
- Rose, D., & Nichols, S. (2013). The lesson of bypassing. *Review of Philosophy and Psychology*, 4(4), 599–619.
- Roskies, A. (2012). Don't panic: Self-authorship without obscure metaphysics. *Philosophical Perspectives*, 26(1), 323–342.
- Searle, J. (1984). *Minds, brains, and science*. Cambridge, MA: Harvard UP.
- Seligman, M., Railton, P., Baumeister, R., & Sripada, C. (2013). Navigating into the future or driven by the past: prospection as an organizing principle of mind. *Perspectives on Psychological Science*, 8(2), 119–141.
- Slooman, S. A. (2005). *Causal models: How people think about the world and its alternatives*. New York: Oxford University Press.
- Smart, J. J. C. (1961). Free will, praise and blame. *Mind*, 70, 291–306.
- Spinoza, B. (1887/1677). *The chief works of Benedict de Spinoza* (Vol. II) (R. Elwes, Trans.). London: Bell and Sons.
- Synofzik, M., Vosgerau, G., & Newen, A. (2008). Beyond the comparator model: A multi-factorial two-step account of agency. *Consciousness and Cognition*, 17(1), 219–239.
- Turner, J., & Nahmias, E. (2006). Are the folk agent causationists? *Mind and Language*, 21(5), 597–609.
- Vihvelin, K. (2004). Free will demystified: A dispositional account. *Philosophical Topics*, 32(1–2), 427–450.
- Wegner, D. (2003). *The illusion of conscious will*. Cambridge, MA: MIT Press.
- Weigel, C. (2011). Distance, anger, freedom: An account of the role of abstraction in compatibilist and incompatibilist intuitions. *Philosophical Psychology*, 24(6), 803–823.
- Woodward, J. (2003). *Making things happen: A theory of causal explanation*. New York: Oxford University Press.
- Yablo, S. (2002). De facto dependence. *Journal of Philosophy*, 99(3), 130–148.